图书情报工作
第63卷 第17期 2019年9月

基于在线评论的网络社区信息可信度评价方法研究

■ 国佳 郭勇 沈旺 潘梦雅
吉林大学管理学院 长春 130025

摘要：[**目的/意义**]提出基于在线评论的网络社区信息可信度评价方法，为信息治理提供有效依据。[**方法/过程**]构建基于在线评论的网络社区信息可信度评价指标体系，利用改进 AHP 理论确定指标权重；利用 LSTM 模型对评论情感分类，采用改进的 D-S 证据理论模型融合情感分类数据作为指标量化计算方法。以知乎网络社区为例，从3个角度计算网络信息内容的可信度：经过筛选的具有可信观点评价的在线评论、所有在线评论、调查问卷。[**结果/结论**]实验结果表明，基于可信观点评论的可信度排序与基于调查问卷的可信度排序基本一致，说明利用在线评论对网络信息可信度进行评价具有一定的可行性。

关键词：在线评论 信息可信度 改进 D-S 证据理论模型 网络社区
分类号：G250
DOI：10.13266/j.issn.0252-3116.2019.17.016

随着网络信息爆炸式增长，网络信息成为人们的主要信息来源。在网络数据访问不受限制的条件下，大量未经证实的信息通过微信、微博、问答社区等平台传播，对人们生活的方方面面产生了重要的影响。虚假信息和不可靠信息的传播呈现泛滥的趋势，对网络信息生态环境造成了极大的破坏。人们对于不熟悉的领域，很难辨别信息的真假。有研究者提出利用“信息治理”来解决这一难题，有效信息治理的前提是对虚假信息的识别，从源头控制信息的传播。早期研究主要是基于信息源、信息内容、媒介平台3个维度构建评价指标体系对信息可信度进行评价，其中，对信息内容的可信度评价主要通过修辞、文本分类等方法，其准确性、专业性等需要通过专家判断，无法处理海量的网络信息资源。因此，本文希望从在线评论中获取群体用户的观点，弥补现有方法的不足，构建更加合理、有效的网络社区信息可信度评价方法。

1 相关理论研究

1.1 网络信息可信度评价指标研究

通过对文献进行梳理，网络信息可信度的评价主要包括3个角度：信息源可信度、信息内容可信度及媒介可信度。

信息源可信度主要是在人际之间、组织中或大众媒介语境下研究不同传播者特征对信息处理过程的影响，传播者可以是个人、小组或组织[1]。因此，对信息源可信度的研究通常分为机构和个人2个方面，机构的类型、权威性等是影响机构信息源可信度的重要因素[2]，作者的身份、地位、声誉等是个人信息源可信度评价的关键因素[3]。高明霞等采用信息总数、是否认证、粉丝总数评价了新浪微博用户信息源的可信度[4]。M. Alrubaian 等从专业度和声誉的角度评价 Twitter 社交网络用户的可信度[5]。

信息内容可信度探讨了用户对信息对象本身的可信度感知[6]。C. N. Wathen 等认为对信息内容的可信度评价应包括专业度、时效性、准确性和相关度[7]。R. Li 等认为论证强度和信息质量影响信息可信度[8]。T. Lucassen 等将信息内容可信度分为语义特征和表面特征：语义特征包括准确性、完整性、全面性、中立性等，表面特征包括长度、参考文献、图片、写作风格等[2]。高明霞等采用文本长度、拼写错误、图表、重复标点、情绪词、转贴数、评论数等表面特征评价信息内容的可信度[4]。G. Sarna 等通过判断网络信息内容中是否存在网络欺凌来评价用户的可信度[9]。李保珍等采用用户评价信息作为信息内容评价的先验经验信息，对信息内容可信度进行评价[10]。

媒介可信度关注信息内容的传播渠道，而不是内

作者简介：国佳(ORCID：0000-0002-1758-1345)，讲师，E-mail：guojiajlu1982@163.com；郭勇(ORCID：0000-0001-7888-3294)，本科生；沈旺(ORCID：0000-0002-8933-5653)，副教授；潘梦雅(ORCID：0000-0002-0319-626X)，硕士研究生。

收稿日期：2019-01-08 修回日期：2019-04-01 本文起止页码：137-144 本文责任编辑：易飞

容的发送者[1]。C. N. Wathen 和 J. Burkell 认为网站的界面设计、加载速度、界面吸引力、可用性、可访问性、交互性和灵活性等技术因素会影响用户可信度感知[7]。R. Li 等验证了媒介依赖、交互性、媒介透明性对信息可信度的影响[8]。B. J. Fogg 等采用问卷调查法进行研究，结果表明网站界面设计、信息设计/结构、信息焦点、广告、信息有用性和准确性、网站创建者、客户服务等会影响网站的可信度[11]。

通过以上文献梳理，我们构建了包含 3 个维度、26 个指标的评价指标体系[12]。但是，在对知乎答案的在线评论进行分析的过程中（根据构建的评价指标体系对知乎 5 万余在线评论进行可信度评价指标编码），发现有些指标很少使用，因此删减统计频率小于 0.1% 的指标[12]，最终形成了包含 3 个维度、21 个指标的评价指标体系，具体内容如表 1 所示：

表 1 基于评论的网络社区信息可信度评价指标框架

维度		指标	指标内涵
信源可信度	机构	机构声誉地位	机构的知名度及在行业中的地位
		机构 URL	机构的网址、该网址是否有机构的详细介绍（名称等）、权威作者总数、赞助商等信息
	用户	作者身份	通常指作者的姓名、职称、专业背景和资历
		作者声誉地位	作者的知名度及其行业地位，包括作者创作量、作者评论转发量等
		作者所属单位	作者所在的单位，单位的性质、地位往往影响用户的信赖程度
		作者动机	作者发布信息的目的，要么是讲述一个主题的真相，要么是为了服务自我提出的偏见信息
		专业知识	作者在专业领域内形成的相对稳定的系统化的知识
信息结构和内容可信度		信息格式	信息中的文本格式和媒体形式是否通用，文本格式包括.doc、.txt 等，媒体格式包括视频、音频等
		逻辑性	文字的表达是否清晰、有逻辑
		写作风格	写作风格是否易于接受
		准确性	信息内容是否真实，是否能够被证实
		客观性	信息对事实描述的公正程度
		时效性	信息发布或更新的时间是否满足用户的要求
		及时性	系统能提供用户所需信息的程度或者对于手头的任务，信息充分更新的程度
		完整性	信息的完整程度，是否存在信息遗漏和成分残缺
		相关性	信息内容与用户所需求的目标值的吻合程度，也就是信息内容是否与用户的需要相关联
		覆盖范围	信息内容涉及主题领域及相关领域的范围
		易理解性	信息内容是否易于理解
		合理性	从自然角度来看，信息描述是否符合自然常识和客观规律；从社会角度来看，信息描述是否符合社会规范和公德伦理
媒介可信度		交互性	媒介信息的发布者与信息用户联系的方式和过程
		权威性	信息的出版和传播渠道是否权威，是否能够令人信服

1.2 网络信息可信度评价方法研究

在网络信息可信度评价方法方面，主要通过抽取可信度指标特征，采用不同的评价模型如 SVM 模型、隐马尔可夫模型、贝叶斯推理理论进行评价。高明霞等提出一个基于信息融合的中文微博可信度评估框架 CCM-IF[4]。G. Sarna 等首先抽取网络欺凌的特征项——链接、坏话、消极/积极情绪、专有名词、代词等，然后利用 SVM 模型评估用户是否可信[9]。M. Park 等利用主题模型和隐马尔科夫模型对在线医疗数据的可信度进行评价[13]。M. Kakol 等构建了一个基于用户评价的网络信息可信度评价预测模型，模型构建的依据是经验数据[14]。Y. Namihira 等提出基于主题和意见分类的自动可信度评价方法[15]。李保珍等提出基于贝叶斯推理理论的网络信息内容可信度的测度模型[10]。孟美任等根据实证结果，选取内容完整性、情感平衡性、评论时效性以及发布者身份明确性 4 类特征，采用 CRFs 模型进行评论可信度 4 级分类，并进行特征组合实验，得到最佳特征组合[16]。

现有信息可信度评价采用的方法如贝叶斯推理理论模型、隐马尔科夫模型、SVM 等都需要经验数据才能构建评价模型，但是网络信息可信度的评价具有很强的主观性，并且现阶段没有可用于网络信息可信度评价的数据集。因此，本文采用加权 D-S 证据理论融合用户评价观点数据，获取各个指标的评分，使用改进 AHP 确定指标权重，最终求出信息可信度评价结果。

2 基于在线评论的网络社区信息可信度评价方法

2.1 在线评论可信度评价情感分类

本文在构建的可信度评价指标体系基础上，对评论进行可信度评价人工标引，通过评论的情感分类获取评论的可信度评价观点数据，如评论"请了解同源重组，想象是靠不住的"是对信息准确性的评价，评论"有道理"是对信息合理性的评价，评论如"看到现在，最理性的答案～"和"我也看出了他在这个地方论证的漏洞"都是对信息合理性进行评价，前者情感分类为积极，说明发表评论的用户认为信息是合理的，后者情感分类为消极，说明发表评论的用户认为信息不合理。评论情感分类的方法主要包括基于词典的方法、支持向量机方法、LDA 模型、机器学习等方法[17-19]。近年来，深度学习在文本情感分类处理研究中取得了不错的成绩，本文采用 LSTM 深度学习模型识别情感极性，处理过程主要包括评论文本数据的预处理、词向量训练、LSTM 情感分类模型构建等。本文对情感分类采用的是两级分类模型，即情感分类结果为积极或者消极，具体分析流程如图 1 所示：

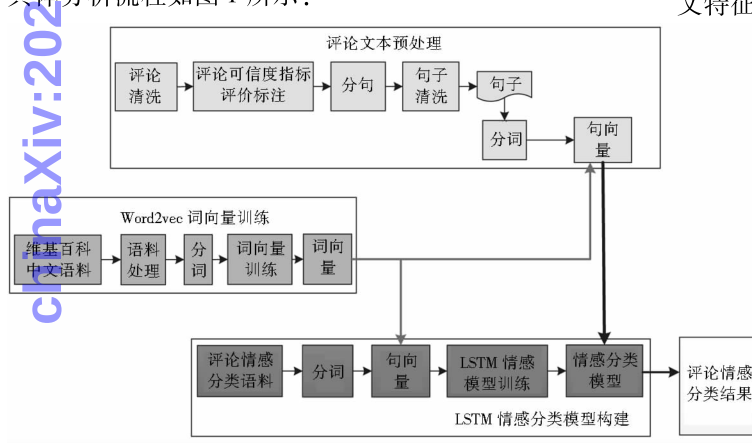


图 1 在线评论可信评价观点识别过程

2.1.1 评论文本数据预处理 包括：①对采集的相关评论数据进行处理，包括冗余数据的清理、不完整数据的删除、符号的转换等数据清洗工作；②根据构建的可信度评价指标体系对评论进行编码，然后选取有可信度指标评价的评论，如评论"数据说话，我服"是对合理性进行评价，是具有可信度指标评价的评论；评论"我有地贫的基因是不是不能献血的?"没有对信息可信度指标进行评价；③对评论文本进行分句，对分句后的文本再次进行清理，删除没有意义的数字、符号等内容；④分词。

2.1.2 基于 LSTM 的情感分析模型构建

（1）word2vec 词向量训练。在 LSTM 情感分析模型构建之前，需要将词转化为词向量。word2vec 模型是 Mikolov 等开发的词向量建模工具[20]，可以用于自然语言问题的解决，在中文评论情感分类应用中取得了较好的效果[21]。由于大规模、高质量的中文语料库比较少，维基语料库对日常生活和常识性的通用性知识的表示效果都很好，知乎作为一个知识型的问答社区，不仅包含日常生活评论，还包含一些通用性知识，采用维基语料库具有可行性。本文 2018 年 8 月 13 日从网络（https://dumps.wikimedia.org/zhwiki/20190101/）下载中文维基百科语料包，将其进行处理（包括繁简转换等），分词，利用 word2vec 进行词向量训练，得到词向量。

（2）LSTM 情感分类模型训练。在自然语言处理中，通常使用的是循环神经网络（recurrent neural network，RNN）。在深度学习语言处理模型中，句子作为序列化的输入，相邻的字构成词，相邻的词构成短语，相邻的短语构成句子。RNN 能将邻近位置关系进行有效整合，更好地完成语言任务，能够很好地利用上下文特征信息，并保留文本的顺序信息，自动选择特征，进行分类。因此，本文采用 LSTM 模型作为情感极性识别的学习模型。具体处理过程如图 1 所示，将获取的情感标注语料进行处理，再进行分词，依据 word2vec 训练出来的词向量模型构建句向量，将句向量输入 LSTM 模型训练，获取 LSTM 情感分类模型。最后，利用构建好的 LSTM 情感分类模型对评论文本内容进行情感分类。

2.2 网络社区信息可信度评价方法

2.2.1 评论融合方法——改进的 D-S 证据理论模型 用户评论数据具有较强的主观性，造成了对信息可信度评价的不确定性。因此，如何运用用户评论数据，并建立合理的评价模型至关重要。证据理论建模工具有很好的证据聚合作用，在专家系统、决策分析、故障诊断、目标识别等信息融合中得到了广泛的应用[22]。在评论融合过程中，出现了 Zadeh 悖论问题，针对该问题，采用改进的 D-S 证据理论模型处理用户评论数据。在用户识别方面，本文将撰写评论用户和点赞该评论用户分别计算，即点赞该评论用户作为与撰写评论观点用户持有相同观点的其他用户。以下是基于改进的 D-S 证据理论的评论融合数学模型。

设 Θ 是一个识别框架。在识别框架 上的基本概率分配是一个 2Θ→[0,1]的函数 m,称为 mass 函数,并且满足 $m(\emptyset)=0$ 且 $\sum_{A\subseteq\Theta} m(A)=1$。此时,m 为框架 Θ 上的基本可信度分配,于是由 $Bel(A)=\sum^{R\subseteq A} m(B)(\forall\subset\Theta)$所定义的函数 $Bel:2^{\Theta}\to[0,1]$为 Θ 上的信度函数。设 $pl:2^{\Theta}\to[0,1]$,$pl(A)=1-Bel(A)$,则称 pl 为 Bel 的似然度函数。该函数表明我们不怀疑 A 的程度或者发现 A 可靠或似真的程度。根据上式,则有 $Pl(A)=\sum_{B\cap A\neq\emptyset} m(B)\ \forall\subset\Theta$。对于 A ,上的有限个 mass 函数 $m_1,m_2,\cdots,m_n$ 的 Dempster 合成规则为:

$$(m_1\oplus m_2\oplus\cdots\oplus m_1)(A)=\frac{1}{K}\sum_{A_1\cap A_2\cap\cdots\cap A_n=A} m_1(A_1)\times m_2(A_2)\times\cdots m_n(A_n)\qquad 式(1)$$

其中,K 为归一化常数,

$$K=\sum_{A_1\cap\cdots\cap A_n\neq\emptyset} m_1(A_1)\cdot m_2(A_2)\cdots m_n(A_n)=1-\sum_{A_1\cap\cdots\cap A_n\neq\emptyset} m_1(A_1)\cdot m_2(A_2)\cdots m_n(A_n)\qquad 式(2)$$

在以上理论基础上,定义评价框架:

(1)设 Θ 是一个识别框架,包含两个不相容的假设命题,其幂集 $2^{\Theta}=\{A_1,A_2\}$。焦元分别是 A_1,A_2,A_1 为积极观点,A_2 为消极观点。

(2)每一个用户的评论作为一项证据,n 个证据 $E_1,E_2,\cdots,E_n$ 的基本信任分配函数为 $m_1,m_2,\cdots,m_n$,评论积极观点概率为 $p_1,p_2,\cdots,p_n$,当 p 值为 1 时,评论为完全积极评论;当 p 值为 0 时,评论为完全消极评论。

(3)根据 Dempster 合成规则,第 i 个特征值对应第 i 个 mass 函数 mi,n 维特征向量 F_n 对应 n 个 mass 函数。当集合 $A=\{A_1\}$时,即观点为积极观点时,对应 n 个 mass 函数$(m_1,m_2,\cdots,m_n)=(p_1,p_2,\cdots,p_n)$,同理,当集合 $A=\{A_2\}$时,对应的 n 个 mass 函数$(m_1,m_2,\cdots,m_n)=(1-p_1,1-p_2,\cdots,1-p_n)$。

(4)设用户积极观点句的数量为 Pos,消极观点句的数量为 Neg,则有:

$$p_n=\frac{Pos_n}{Pos_n+Neg_n}\qquad 式(3)$$

(5)将评论信息进行融合。由于评论的情感极性矛盾较大,导致证据理论合成悖论,本文采用加权平均法解决这一问题[23]。具体方法如下:

步骤一:计算证据间的相似系数,列出相似矩阵。证据 E1 和 E2 间的相似系数可以表示为:

$$d_{12}=\frac{\sum_{A_i\cap B_j=A_k\neq\emptyset} m_1(A_1)m_2(A_2)}{\sqrt{(\sum m_1^2(A_1))(m_2^2(A_2))}}\quad d_{12}\in[0,1]\qquad 式(4)$$

根据用式(4)计算出证据之间的相似系数,以矩阵的形式表示:

$$s=\begin{matrix} 1 & d_{12} & \cdots & d_{1n} \\ d_{21} & 1 & \cdots & d_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ d_{n1} & d_{n2} & \cdots & 1 \end{matrix}\qquad 式(5)$$

步骤二:求出各证据的支持度。将矩阵的每行相加得到每个证据对的 Ei 支持度为:

$$aup(m_i)=\sum_{j=1}^{n} d_{ij}(i,j=12,\cdots,n)\qquad 式(6)$$

步骤三:求出各证据的可信度。将证据的支持度归一化可得到证据 Ei 的可信度:

$$Crd(m_i)=\frac{sup(m_i)}{\sum_{i=1}^{n} sup(m_i)}(i,j=1,2,\cdots,n)\qquad 式(7)$$

步骤四:把可信度作为权重,对证据的基本信任分配进行加权平均。

步骤五:用 D-S 合成规则式(1)、(2)合成加权平均证据,取加权平均证据结果作为各个指标的可信度评价结果。

2.2.2 指标权重确定方法 本文采用改进的 AHP 方法确定指标的权重。传统的 AHP 方法需要根据一致性检验结果判断结果的有效性,并且需要多次调整才能通过一致性检验,本文借鉴梁樑等[24]的研究,确定指标权重,该方法采用最优传递矩阵一次性确定权重,不需要进行一致性检验。其计算过程如下:

步骤一:使用德尔菲法构造判别矩阵 A,对体系中的各级指标进行两两比较,确定各个指标的相对重要性,各级指标标度如表 2 所示:

表 2 各级指标矩阵比较标度

重要程度	说明	标度 f(x, y)
属性 x 较 y 同等重要	x、y 对总目标的贡献相同	1
属性 x 较 y 稍微重要	x 的贡献稍大于 y	3
属性 x 较 y 明显重要	x 的贡献明显大于 y	5
属性 x 较 y 特别重要	x 的贡献特别明显大于 y	7
属性 x 较 y 极端重要	x 的贡献以压倒优势大于 y	9
属性 x 较 y 介于各等级之间	相邻两判断的折中	2,4,6,8

步骤二:计算判别矩阵 A 的最优传递矩阵:

$$b_{ij}=\lg a_{ij}(i,j=1,2,\cdots,n)\qquad 式(8)$$

$$c_{ij}=\frac{1}{n}\sum_{k=1}^{n}(b_{ik}-b_{jk})(1,j=1,2,\cdots,n)\qquad 式(9)$$

步骤三:计算拟优矩一致矩阵:

$$a_{ij}^{*}=10^{c_{ij}}(i,j=1,2,\cdots,n)\qquad 式(10)$$

步骤四:求 A^{*} 特征向量(方根法)得:

$$w_i^* = (\bar{w}_1, \bar{w}_2, \cdots, \bar{w}_n)^T (i = 1,2,\cdots,n) \quad 式(11)$$

步骤五：归一化得到指标权重：

$$w_i = \frac{\bar{w}_i}{\sum_{i=1}^{n} \bar{w}_i} = (w_1, w_2, \cdots, w_n)^T (i = 1,2,\cdots. n) \quad 式(12)$$

3 应用研究

3.1 数据采集及预处理

实证数据来源于知乎问答社区，在对知乎的话题进行分析后，我们选择用户关注和争论较多的献血和转基因话题作为采集主题。利用网络爬虫，在 2017 年 11 月 15－23 日爬取了54 021条评论及相关信息（包括问题、答案、点赞数等）。我们采用扎根理论对 54 021 条评论进行了信息可信度评价编码，对可信度进行评价的评论共有 29 096 条。由于答案较多，我们选择了可信度观点评论较多的 10 个答案作为样本。每个答案的基本信息如表 3 所示：

表 3 知乎答案基本信息表

答案序号	话题	观点	信息内容	作者身份	知乎身份	评论数量（条）	可信观点评论数量（条）
1	献血	支持	鼓励献血，献血的害处不太大，但绝非有益，鼓励在身体健康情况下献血	临床医学博士	无	611	231
2	献血	不支持	不鼓励献血，在学校献血车献血过程感觉不好	土木工程专业学生	无	456	159
3	献血	反对	反对献血，认为免费献血过程中存在黑幕	匿名用户	无	713	494
4	献血	支持	鼓励献血，说明血站的困难和可能存在的问题	医学行业人员	医学话题优秀回答者	998	212
5	转基因	中立	对支持转基因者表示忧虑，不赞同反对转基因者	转基因行业人员	无	452	325
6	转基因	支持	支持，提供证明转基因安全的证据图片	无	无	453	218
7	转基因	支持	支持转基因，认为反转基因是“穷人思维”	无	无	539	366
8	转基因	反对	反转基因	知名网络小说作者	无	292	231
9	转基因	中立	中立，转基因技术可以带来益处，但是也存在风险	律师	无	236	188
10	转基因	支持	支持，支持转基因商业化	保险行业人员	无	249	155

在对评论进行分句后，本文对无情感含义的字符串、数字进行了删除，获得 15 562 条句子，之后采用结巴分词工具对评论句子进行分词。

3.2 在线评论可信度评价情感分类

首先，从 Wiki 百科下载中文最新数据作为词向量训练语料，去除帮助页面和重定向等无用页面。其次，过滤特殊的非文本标记。最后对于文本信息进行繁体到简体的转化。清洗语料，去除标点、数字、非中文字符等。在对语料进行分词后进行词向量训练，词向量的维度为 300 维，滑动窗口大小为 5。训练后获取词向量模型库。

本文通过分析采集的评论，构建了包含 5 000 条人工标引情感极性的网络社区评论情感分类语料库。使用结巴软件对语料进行分词，在词向量基础上构建句向量。LSTM 模型输入节点设置为 256，隐藏节点设置为 128。经过训练获得 LSTM 情感分类模型。模型分类的准确率为 90.01%。

将经过清洗和处理的在线评论转化为句向量，输入到 LSTM 情感分类模型中，获得基于 LSTM 模型的在线评论的情感分类。

3.3 基于在线评论的网络社区信息可信度评价结果

依据 2.2.1 的评论融合方法，对情感分类数据进行融合计算。在计算过程中，不只计算评论用户的观点，还计算点赞用户的观点，即当用户对某一评论点赞，则认为他具有相同观点，并作为一个独立的证据进行计算。融合各个用户的评论情感证据，得到每个答案的可信度指标值见表 4。

依据 2.2.2 的权重计算方法，使用专家判断矩阵计算指标权重，指标权重见表 5。最后，计算出每个问题的信息可信度评价结果见表 6。

3.4 评价结果分析

信息可信度是用户根据经验判断事情为真的相信程度，是主观性的感知，知乎社区中答案的赞同数、评论数等并不能反映出答案的可信程度（例如，只包含一句搞笑的话的答案得到了更高的赞同，并不是因为用户相信这条答案，而是由于用户觉得这个答案有趣）。因此，为了验证评价结果的有效性，本文选择了 50 位调查者对每个问题的可信度进行了问卷调查，去除其中的无效问卷 6 份，最终采集问卷 44 份，管中窥豹，希望能够获得用户对答案的可信度主观判断。调查者的年龄范围在 18－35 岁，涉及 19 个行业或专业，学历为高中、本科和硕士。调查者需要详细查看每个答案，并填写调查问卷。调查问题一是问题的可信度为可信、不确定及不可信，对应数值为 0，0.5 和 1；二是问题的

表 4 基于可信观点评论情感分析的信息可信度评价指标结果

问题编号	机构声誉地位	机构URL	作者身份	作者声誉地位	作者所属单位	作者动机	专业知识	信息格式	逻辑性	写作风格	准确性	客观性	时效性	及时性	完整性	相关性	覆盖范围	易理解性	合理性	交互性	权威性
1	1	1	1	1	1	1	1	1	1	1	0.062 6	1	1	0.5	0.070 7	1	1	1	0.095 0	1	1
2	0	0	0	1	1	0	0.020 1	1	0	0	0	0.057 8	1	0	0.070 7	0.070 7	0.666 7	0	0	1	0
3	1	1	1	1	1	0	0.100 7	1	0.051 2	1	0.093 4	0.157 7	1	1	0	0	1	0	0.127 3	1	0
4	1	1	1	1	1	1	1	1	0.079 9	1	0.170 1	1	1	1	1	1	0.147 0	0.079 9	0		1
5	1	1	1	1	1	0	0.008 0	1	0.158 0	1	0.069 8	0.064 9	0	0	0.342 3	1	1	0.141 1	0.104 9	1	1
6	1	1	0.132 7	1	1	1	0.036 0	1	0.154 4	1	0.129 6	0.088 1	1	1	0	1	1	0.092 6	0.655 1	1	1
7	0	0	0	0	0	0.617 7	0.280 0	0	0.771 1	0.433 7	0.246 1	0.269 9	0	0	0	0	1	1	0.281 3	0	0
8	1	1	0	1	1	0	0.061 6	1	0.092 8	1	0.093 9	0.074 9	0.224 4	1	1	0.104 3	0.5	1	0.058 8	1	1
9	1	1	0.528 3	1	1	1	0.680 8	1	0	1	0.403 1	0.678 4	1	0.5	0	1	0	0.408 3	0.603 7	0	0
10	1	1	0.246 1	0.269 9	1	0.617 7	0.280 0	1	1.280 0	0.433 7		1	1	1	1	0	1	1	0.281 3	1	1

表 5 各指标权重

指标	机构声誉地位	机构URL	作者身份	作者声誉地位	作者所属单位	作者动机	专业知识	信息格式	逻辑性	写作风格	准确性	客观性	时效性	及时性	完整性	相关性	覆盖范围	易理解性	合理性	交互性	权威性
权重	0.038 9	0.008 2	0.025 0	0.034 0	0.027 7	0.075 1	0.069 4	0.007 2	0.085 9	0.007 6	0.127 9	0.104 6	0.008 0	0.008 0	0.015 9	0.171 5	0.017 2	0.014 1	0.109 3	0.008 1	0.036 3

表 6 基于调查问卷可信度评价结果

答案序号	基于调查问卷的可信度值	基于可信观点评论情感分析的信息可信度值	基于所有评论情感分析的信息可信度值
1	0.847 8	0.762 4	0.304 3
2	0.630 4	0.117 2	0.594 4
3	0.391 3	0.243 8	0.094 7
4	0.782 6	0.669 8	0.251 8
5	0.456 5	0.430 5	0.049 2
6	0.565 2	0.565 6	0.360 2
7	0.434 8	0.257 2	0.071 7
8	0.456 5	0.272 8	0.094 3
9	0.652 2	0.637 0	0.186 2
10	0.543 5	0.519 4	0.113 7

熟悉度，对应数值为 1，3，5。将问题的熟悉度作为权重，计算每个答案的可信度，具体计算方法如见式(13)，计算结果见表 6。同时，为了验证使用所有评论数据是否能够对信息内容可信度评价，本文将所有评论进行情感分析，并采用证据融合方法进行计算，获得基于所有评论的网络信息内容可信度值，见表 6。

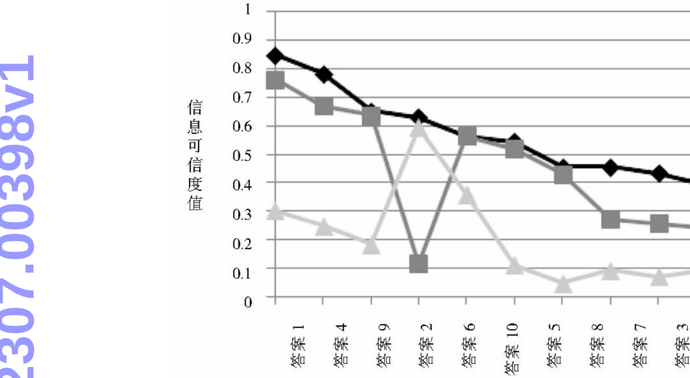
$$Cr = c_i \sum_{i=1}^{n} \frac{f_i}{(f_1 + f_2 + \cdots + f_n)} (i = 1, 2, \cdots, n) \quad 式(13)$$

其中，C_r 为答案的可信度评价最终结果，共有 n 个用户对答案可信度进行评价，f_i 为第 i 个用户对问题的熟悉程度，c_i 为第 i 个用户对问题的可信度评价数值。本文对 3 种方法获得的评价结果按照数值大小进行了重新排列，以便于结果分析，见表 7。通过对可信度值的排列可以发现，在去掉答案 2 后，基于调查问卷的可信度值排序和基于可信观点评论情感分析的可信度值排序基本一致，说明了基于可信观点情感分析的可信度评价方法的有效性。基于所有评论情感分析的可信度评价结果与基于问卷的可信度评价结果不一致，说明不能用所有评论情感分析对信息内容的可信度进行评价。基于调查问卷的可信度值稍高于基于可信观点评论情感分析的可信度值，如图 2 所示，这是由于用户对信息内容产生质疑后，更容易导致可信度观点评论。。从图 2 中我们还观察到了一个有趣的现象，当可信度值的区间为 0.4－0.7 时，基于问卷的可信度值与基于可信观点评论情感分析的可信度值基本一致，说明当对可信度评价不确定时，基于可信观点评论情感分析的可信度评价更接近实际评价，当可信度评价结果趋于可信或不可信时，基于可信观点评论情感分析的可信度评价值低于实际值。

本文对异常点进行了分析，调查者对答案 2 的可信度评价较高，排名第 4，而基于评论情感分析的可信度值排在第 10 位，差别较大。答案 2 的观点为“不鼓励献血，在学校献血车献血过程感觉不好”，评论内容多为对自己献血的事实描述，如“我第一次献血的时候没吃饭还晕了”“从那以后身体就不如以前了”“刚献完直接晕那五分钟”等，根据情感分析结果，这些评论都是消极评价，代表答案内容不可信，但是实际上其观点与答案观点一致，说明答案是可信的，因此，导致了评价结果差异。该问题为进一步研究提出了方向，即

表 7　基于调查问卷的可信度值与基于评论情感分析的可信度值排序比较

排名	序号	基于调查问卷可信度值	序号	基于可信观点评论情感分析的可信度值	序号	基于所有评论情感分析的可信度值
1	1	0.847 8	1	0.762 4	2	0.594 4
2	4	0.782 6	4	0.669 8	6	0.360 2
3	9	0.652 2	9	0.637 0	1	0.304 3
4	2	0.630 4	6	0.565 6	4	0.251 8
5	6	0.565 2	10	0.519 4	9	0.186 2
6	10	0.543 5	5	0.430 5	10	0.113 7
7	5	0.456 5	8	0.272 8	3	0.094 7
8	8	0.456 5	7	0.257 2	8	0.094 3
9	7	0.434 8	3	0.243 8	7	0.071 7
10	3	0.391 3	2	0.117 2	5	0.049 2



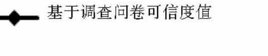


图 2　3 种评价结果比较

在当评论内容为事实描述的时候，情感分类不能代表评论的可信评价，需要通过验证事实描述观点与答案内容观点的一致性评价信息的可信度。

4　结论

本文提出一种基于在线评论计算网络信息的可信度的评价方法，首先构建了基于在线评论的网络信息可信度评价指标体系，利用改进的 AHP 模型确定权重；采用 LSTM 模型计算评论的情感极性，通过改进的证据理论模型对用户评论数据进行融合，结合权重最终计算出网络信息的可信度值。通过对知乎两个主题下 10 个样本的评论信息的处理，对提出的方法进行试验验证。试验结果表明在去除一个异常值后，基于问卷的网络信息可信度值与基于可信观点评论的网络可信度值排序基本一致，可以有效对网络信息的可信度进行排序；使用所有评论对网络信息进行评价的准确率较低，基于评论对网络信息可信度进行评价前必须进行可信观点评论筛选。同时本文也存在一些局限，样本类型只选择了知乎两个主题的评论，评论数量规模有限，还有异常值的出现，说明网络信息可信度评价是一个复杂的过程，当评论包含对事实描述的时候，需要进一步研究评论观点与信息内容观点的一致性，后续研究希望扩大样本量和范围，综合利用多种手段解决这些问题。

参考文献：

[1] KIOUSIS S. Public trust or mistrust? perceptions of media credibility in the information age[J]. Mass communication & society, 2001, 4(4):381－403.
[2] LUCASSEN T, SCHRAAGEN J M. Factual accuracy and trust in information: the role of expertise[J]. Journal of the American Society for Information Science and Technology, 2011,62(7):1232－1242
[3] WESTERMAN D, SPENCE P R, HEIDE B V D. Social media as information source: recency of updates and credibility of information [J]. Journal of computer-mediated communication, 2014, 19(2):171－183.
[4] 高明霞，陈福荣. 基于信息融合的中文微博可信度评估方法[J]. 计算机应用, 2016, 36(8):2071－2075.
[5] ALRUBAIAN M, AL-QURISHI M, AL-RAKHAMI M, et al. Reputation-based credibility analysis of twitter social network users [J]. Concurrency & computation practice & experience, 2017, 29(7):1－12.

[6] RIEH S Y, DANIELSON D R. Credibility: a multidisciplinary framework [J]. Annual review of information science & technology, 2010, 41(1):307－364.
[7] WATHEN C N, BURKELL J. Believe it or not: factors influencing credibility on the Web[J]. Journal of the Association for Information Science & Technology, 2002, 53(2):134－144.
[8] LI R, SU A. Factors influencing information credibility on social media platforms: evidence from facebook pages[J]. Procedia computer science, 2015, 72:314－328.
[9] SARNA G, BHATIA M P S. Content based approach to find the credibility of user in social networks: an application of cyberbullying[J]. International journal of machine learning & cybernetics, 2015, 8(2):1－13.
[10] 李保珍, 王亚, 周可. 基于贝叶斯理论的社会化媒体网络信息内容可信度测度[J]. 数据分析与知识发现, 2017, 1(6):83－92.
[11] FOGG B J, SOOHOO C, DANIELSON D R, et al. How do users evaluate the credibility of Websites?: a study with over 2,500 participants[M]. New York: Association for Computing Machinery, 2003.
[12] 沈旺,康霄普,王佳馨,等. 用户视角下社会化问答社区信息可信度评价研究[J]. 图书情报工作,2018,62(17):104－111.
[13] PARK M, SAMPATHKUMAR H, LUO B, et al. Content-based assessment of the credibility of online healthcare information[C]// IEEE international conference on big data. USA:IEEE,2013:51－58.
[14] KAKOL M, NIELEK R, WIERZBICKI A. Understanding and predicting Web content credibility using the content cedibility corpus [J]. Information processing & management, 2017, 53(5):1043－1061.
[15] NAMIHIRA Y, SEGAWA N, IKEGAMI Y, et al. High precision credibility analysis of information on Twitter[C]// International conference on signal-image technology & internet-based systems. Kyoto:IEEE, 2013:909－915.
[16] 孟美任, 丁晟春. 在线中文商品评论可信度研究[J]. 现代图书情报技术, 2013, 29(9):60－66.
[17] 唐晓波,兰玉婷. 基于特征本体的微博产品评论情感分析[J]. 图书情报工作,2016,60(16):121－127,136.
[18] 曹宇,王名扬,贺惠新. 情感词典扩充的微博文本多元情感分类研究[J]. 情报杂志,2016,35(10):185－189.
[19] 何跃,赵书朋,何黎. 基于情感知识和机器学习算法的组合微文情感倾向分类研究[J]. 情报杂志,2018,37(5):189－194.
[20] 王勤勤, 张玉红, 李培培,等. 基于 word2vec 的跨领域情感分类方法[J]. 计算机应用研究, 2018, 35(10):50－53.
[21] 张冬雯,杨鹏飞,许云峰. 基于 word2vec 和 SVMperf 的中文评论情感分类研究[J]. 计算机科学,2016,43(S1):418－421,447.
[22] 何晶晶,蔡德胜,介飞,等. 利用 D-S 证据理论进行特征融合的同义实体识别[J]. 计算机应用研究,2018,35(5):1429－1433.
[23] 杨风暴, 王肖霞. D-S 证据理论的冲突证据合成方法[M]. 北京:国防工业出版社, 2010.
[24] 梁樑, 盛昭翰, 徐南荣. 一种改进的层次分析法[J]. 系统工程, 1989(3):5－7.

作者贡献说明:
国佳: 论文撰写、数据分析及论文终稿修订;
郭勇: 数据分析和处理;
沈旺: 研究思路设计;
潘梦雅: 数据分析。

Credibility Evaluating Method of Online Community Information Based on Online Reviews

Guo Jia Guo Yong Shen Wang Pan Mengya
School of Management, Jilin University, Changchun 130025

Abstract: [**Purpose/significance**] A social Q&A community information credibility evaluation method based on online reviews was proposed, aiming to provid an effective methods for information governing. [**Method/process**] The paper constructed the evaluation system of online community information reliability based on online reviews, and the indicator weight was determined by improved AHP theory. The LSTM model was used to classify the reviews emotion, and the improved D-S evidence theory model was used to fuse the emotion classification data. Taking ZhiHu Community as an example, the credibility of network information content was calculated from three perspectives: screened online reviews with credible opinion evaluation, all online reviews and questionnaires. [**Result/conclusion**] The experimental results showed that the ranking of the credibility value of this method was basically the same as that of the ranking obtained by the questionnaire. It showed that it is feasible to evaluate the credibility of the network information by online review.

Keywords: online reviews information credibility modified D-S evidential theory online community